RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _____ 10/560,961 _____
Source: _____ JFWP _____
Date Processed by STIC: _____ 12/30/2005 _____

ENTERED

IFWP

RAW SEQUENCE LISTING   DATE: 12/30/2005
PATENT APPLICATION: US/10/560,961   TIME: 09:43:13

          Input Set : A:\15450PCT00.ST25.txt
          Output Set: N:\CRF4\12302005\J560961.raw

```
      3 <110> APPLICANT: Pharmexa A/S
      5 <120> TITLE OF INVENTION: PURIFICATION OF HER-2 VARIANTS
      7 <130> FILE REFERENCE: P1020DK00
C-->  9 <140> CURRENT APPLICATION NUMBER: US/10/560,961
C-->  9 <141> CURRENT FILING DATE: 2005-12-14
      9 <160> NUMBER OF SEQ ID NOS: 2
     11 <170> SOFTWARE: PatentIn version 3.2
     13 <210> SEQ ID NO: 1
     14 <211> LENGTH: 5661
     15 <212> TYPE: DNA
     16 <213> ORGANISM: Artificial sequence
     18 <220> FEATURE:
     19 <223> OTHER INFORMATION: Recombinant expression plasmid derived from pMT
     22 <220> FEATURE:
     23 <221> NAME/KEY: polyA_signal
     24 <222> LOCATION: (263)..(268)
     25 <223> OTHER INFORMATION: SV40 late polyadenylation site
     27 <220> FEATURE:
     28 <221> NAME/KEY: misc_feature
     29 <222> LOCATION: (1579)..(2439)
     30 <223> OTHER INFORMATION: Ampicillin resistance gene, encoded by complementary strand
     32 <220> FEATURE:
     33 <221> NAME/KEY: promoter
     34 <222> LOCATION: (3050)..(3415)
     35 <223> OTHER INFORMATION: Metallothionein promoter
     37 <220> FEATURE:
     38 <221> NAME/KEY: RBS
     39 <222> LOCATION: (3493)..(3501)
     40 <223> OTHER INFORMATION: Kozak-like sequence
     42 <220> FEATURE:
     43 <221> NAME/KEY: CDS
     44 <222> LOCATION: (3502)..(5592)
     45 <223> OTHER INFORMATION: DNA encoding immunogenic, his-tagged variant of human HER-2
     47 <220> FEATURE:
     48 <221> NAME/KEY: sig_peptide
     49 <222> LOCATION: (3502)..(3555)
     50 <223> OTHER INFORMATION: BiP signal sequence
     52 <220> FEATURE:
     53 <221> NAME/KEY: misc_feature
     54 <222> LOCATION: (3556)..(3597)
     55 <223> OTHER INFORMATION: Histidine tag
     57 <220> FEATURE:
     58 <221> NAME/KEY: mat_peptide
```

```
59 <222> LOCATION: (3556)..(5589)
61 <220> FEATURE:
62 <221> NAME/KEY: misc_feature
63 <222> LOCATION: (3598)..(3603)
64 <223> OTHER INFORMATION: Dipeptidase stop sequence
66 <220> FEATURE:
67 <221> NAME/KEY: misc_feature
68 <222> LOCATION: (3604)..(5589)
69 <223> OTHER INFORMATION: Gene coding for the hHER2MA5-5DUH protein
71 <220> FEATURE:
72 <221> NAME/KEY: misc_feature
73 <222> LOCATION: (4357)..(4401)
74 <223> OTHER INFORMATION: Diphtheria toxoid P2 epitope
76 <220> FEATURE:
77 <221> NAME/KEY: misc_feature
78 <222> LOCATION: (5500)..(5562)
79 <223> OTHER INFORMATION: Diphtheria toxoid P30 epitope
81 <400> SEQUENCE: 1
82 ggccgctcga gtctagaggg cccttcgaag gtaagcctat ccctaaccct ctcctcggtc     60
84 tcgattctac gcgtaccggt catcatcacc atcaccattg agtttaaacc cgctgatcag    120
86 cctcgactgt gccttctaag atccagacat gataagatac attgatgagt ttggacaaac    180
88 cacaactaga atgcagtgaa aaaaatgctt tatttgtgaa atttgtgatg ctattgcttt    240
90 atttgtaacc attataagct gcaataaaca agttaacaac aacaattgca ttcattttat    300
92 gtttcaggtt cagggggagg tgtgggaggt tttttaaagc aagtaaaacc tctacaaatg    360
94 tggtatggct gattatgatc agtcgacctg caggcatgca agcttggcgt aatcatggtc    420
96 atagctgttt cctgtgtgaa attgttatcc gctcacaatt ccacacaaca tacgagccgg    480
98 aagcataaag tgtaaagcct ggggtgccta atgagtgagc taactcacat taattgcgtt    540
100 gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc cagctgcatt aatgaatcgg    600
102 ccaacgcgcg gggagaggcg gtttgcgtat tgggcgctct tccgcttcct cgctcactga    660
104 ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat    720
106 acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa aaggccagca    780
108 aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgcccccc    840
110 tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata    900
112 aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc    960
114 gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcatagctc   1020
116 acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga   1080
118 accccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc   1140
120 ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag   1200
122 gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag   1260
124 gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag   1320
126 ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt gcaagcagca   1380
128 gattacgcgc agaaaaaaag gatctcaaga agatccttttg atcttttcta cggggtctga   1440
130 cgctcagtgg aacgaaaact cacgttaagg attttggtc atgagattat caaaaaggat   1500
132 cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga   1560
134 gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg   1620
136 tctatttcgt tcatccatag ttgcctgact ccccgtcgtg tagataacta cgatacggga   1680
138 gggcttacca tctggcccca gtgctgcaat gataccgcga gacccacgct caccggctcc   1740
140 agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg gtcctgcaac   1800
142 tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc   1860
```

```
144 agttaatagt ttgcgcaacg ttgttgccat tgctacaggc atcgtggtgt cacgctcgtc   1920
146 gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc   1980
148 catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca gaagtaagtt   2040
150 ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc   2100
152 atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagtg   2160
154 tatgcggcga ccgagttgct cttgcccggc gtcaatacgg gataataccg cgccacatag   2220
156 cagaacttta aaagtgctca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat   2280
158 cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc   2340
160 atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa   2400
162 aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta   2460
164 ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa   2520
166 aaataaacaa ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga   2580
168 aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgttc   2640
170 gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc agctcccgga cacggtcaca   2700
172 gcttgtctgt aagcggatgc cgggagcaga caagcccgtc agggcgcgtc agcgggtgtt   2760
174 ggcgggtgtc ggggctggct taactatgcg gcatcagagc agattgtact gagagtgcac   2820
176 catatgcggt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat caggcgccat   2880
178 tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta   2940
180 cgccagctgg cgaaggggg atgtgctgca aggcgattaa gttgggtaac gccagggttt   3000
182 tcccagtcac gacgttgtaa aacgacggcc agtgccagtg aattaattcg ttgcaggaca   3060
184 ggatgtggtg cccgatgtga ctagctcttt gctgcaggcc gtcctatcct ctggttccga   3120
186 taagagaccc agaactccgg cccccaccg cccaccgcca cccccataca tatgtggtac   3180
188 gcaagtaaga gtgcctgcgc atgccccatg tgccccacca agagttttgc atcccataca   3240
190 agtccccaaa gtggagaacc gaaccaattc ttcgcgggca gaacaaaagc ttctgcacac   3300
192 gtctccactc gaatttggag ccggccggcg tgtgcaaaag aggtgaatcg aacgaaagac   3360
194 ccgtgtgtaa agccgcgttt ccaaaatgta taaaaccgag agcatctggc caatgtgcat   3420
196 cagttgtggt cagcagcaaa atcaagtgaa tcatctcagt gcaactaaag gggggatcta   3480
198 gatcggggta ccaaagtcac c atg aag ttg tgc atc ttg ctg gcc gtc gtg   3531
199                         Met Lys Leu Cys Ile Leu Leu Ala Val Val
200                                 -15                     -10
202 gcc ttc gtg ggc ctg tcg ctg ggc atg aag cac caa cac caa cat caa   3579
203 Ala Phe Val Gly Leu Ser Leu Gly Met Lys His Gln His Gln His Gln
204             -5              -1  1                 5
206 cat caa cat caa cat caa gcc ccc tcc acc caa gtg tgt acc ggc aca   3627
207 His Gln His Gln His Gln Ala Pro Ser Thr Gln Val Cys Thr Gly Thr
208     10                  15                  20
210 gac atg aag ctg cgg ctc cct gcc agt ccc gag acc cac ctg gac atg   3675
211 Asp Met Lys Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met
212 25                  30                  35                  40
214 ctc cgc cac ctc tac cag ggc tgc cag gtg gtg cag gga aac ctg gaa   3723
215 Leu Arg His Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu
216             45                  50                  55
218 ctc acc tac ctg ccc acc aat gcc agc tta agt ttc ctg cag gat atc   3771
219 Leu Thr Tyr Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile
220             60                  65                  70
222 cag gag gtg cag ggc tac gtg ctc atc gct cac aac caa gtg agg cag   3819
223 Gln Glu Val Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln
224         75                  80                  85
226 gtc cca ctg cag agg ctg cgg att gtg cga ggc acc cag ctc ttt gag   3867
```

```
227 Val Pro Leu Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu
228     90                  95                  100
230 gac aac tat gcc ctg gcc gtg cta gac aat gga gac ccg ctg aac aat    3915
231 Asp Asn Tyr Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn
232 105                 110                 115                 120
234 acc acc cct gtc aca ggg gcc tcc cca gga ggc ctg cgg gag ctg cag    3963
235 Thr Thr Pro Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln
236                 125                 130                 135
238 ctt cga agc ctc aca gag atc ttg aaa gga ggg gtc ttg atc cag cgg    4011
239 Leu Arg Ser Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg
240                 140                 145                 150
242 aac ccc cag ctc tgc tac cag gac acg att ttg tgg aag gac atc ttc    4059
243 Asn Pro Gln Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe
244     155                 160                 165
246 cac aag aac aac cag ctg gct ctc aca ctg ata gac acc aac cgc tct    4107
247 His Lys Asn Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser
248     170                 175                 180
250 cgg gcc tgc cac ccc tgt tct ccg atg tgt aag ggc tcc cgc tgc tgg    4155
251 Arg Ala Cys His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp
252 185                 190                 195                 200
254 gga gag agt tct gag gat tgt cag agc ctg acg cgc act gtc tgt gcc    4203
255 Gly Glu Ser Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala
256                 205                 210                 215
258 ggt ggc tgt gcc cgc tgc aag ggg cca ctg ccc act gac tgc tgc cat    4251
259 Gly Gly Cys Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His
260                 220                 225                 230
262 gag cag tgt gct gcc ggc tgc acg ggc ccc aag cac tct gac tgc ctg    4299
263 Glu Gln Cys Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu
264     235                 240                 245
266 gcc tgc ctc cac ttc aac cac agt ggc atc tgt gag ctg cac tgc cca    4347
267 Ala Cys Leu His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro
268     250                 255                 260
270 gcc ctg gtc cag tac atc aaa gct aac tcc aaa ttc atc ggt atc acc    4395
271 Ala Leu Val Gln Tyr Ile Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr
272 265                 270                 275                 280
274 gag ctg cgg tat aca ttc ggc gcc agc tgt gtg act gcc tgt ccc tac    4443
275 Glu Leu Arg Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr
276                 285                 290                 295
278 aac tac ctt tct acg gac gtg gga tcc tgc acc ctc gtc tgc ccc ctg    4491
279 Asn Tyr Leu Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu
280                 300                 305                 310
282 cac aac caa gag gtg aca gca gag gat gga aca cag cgg tgt gag aag    4539
283 His Asn Gln Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys
284                 315                 320                 325
286 tgc agc aag ccc tgt gcc cga gtg tgc tat ggt ctg ggc atg gag cac    4587
287 Cys Ser Lys Pro Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His
288     330                 335                 340
290 ttg cga gag gtg agg gca gtt acc agt gcc aat atc cag gag ttt gct    4635
291 Leu Arg Glu Val Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala
```

RAW SEQUENCE LISTING   DATE: 12/30/2005
PATENT APPLICATION: US/10/560,961   TIME: 09:43:13

Input Set : A:\15450PCT00.ST25.txt
Output Set: N:\CRF4\12302005\J560961.raw

```
292 345                   350                 355                 360
294 ggc tgc aag aag atc ttt ggg agc ctg gca ttt ctg ccg gag agc ttt   4683
295 Gly Cys Lys Lys Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe
296                   365                 370                 375
298 gat ggg gac cca gcc tcc aac act gcc ccg ctc cag cca gag cag ctc   4731
299 Asp Gly Asp Pro Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu
300                   380                 385                 390
302 caa gtg ttt gag act ctg gaa gag atc aca ggt tac cta tac atc tca   4779
303 Gln Val Phe Glu Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser
304             395                 400                 405
306 gca tgg ccg gac agc ctg cct gac ctc agc gtc ttc cag aac ctg caa   4827
307 Ala Trp Pro Asp Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln
308       410                 415                 420
310 gta atc cgg gga cga att ctg cac aat ggc gcc tac tcg ctg acc ctg   4875
311 Val Ile Arg Gly Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu
312 425                   430                 435                 440
314 caa ggg ctg ggc atc agc tgg ctg ggg ctg cgc tca ctg agg gaa ctg   4923
315 Gln Gly Leu Gly Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu
316                   445                 450                 455
318 ggc agt gga ctg gcc ctc atc cac cat aac acc cac ctc tgc ttc gtg   4971
319 Gly Ser Gly Leu Ala Leu Ile His His Asn Thr His Leu Cys Phe Val
320             460                 465                 470
322 cac acg gtg ccc tgg gac cag ctc ttt cgg aac ccg cac caa gct ctg   5019
323 His Thr Val Pro Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu
324       475                 480                 485
326 ctc cac act gcc aac cgg cca gag gac gag tgt gtg ggc gag ggc ctg   5067
327 Leu His Thr Ala Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu
328       490                 495                 500
330 gcc tgc cac cag ctg tgc gcc cga ggg cac tgc tgg ggt cca ggg ccc   5115
331 Ala Cys His Gln Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro
332 505                   510                 515                 520
334 acc cag tgt gtc aac tgc agc cag ttc ctt cgg ggc cag gag tgc gtg   5163
335 Thr Gln Cys Val Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val
336                   525                 530                 535
338 gag gaa tgc cga gta ctg cag ggg ctc ccc agg gag tat gtg aat gcc   5211
339 Glu Glu Cys Arg Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala
340             540                 545                 550
342 agg cac tgt ttg ccg tgc cac cct gag tgt cag ccc cag aat ggc tca   5259
343 Arg His Cys Leu Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser
344       555                 560                 565
346 gtg acc tgt ttt gga ccg gag gct gac cag tgt gtg gcc tgt gcc cac   5307
347 Val Thr Cys Phe Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His
348       570                 575                 580
350 tat aag gac cct ccc ttc tgc gtg gcc cgc tgc ccc agc ggt gtg aaa   5355
351 Tyr Lys Asp Pro Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys
352 585                   590                 595                 600
354 cct gac ctc tcc tac atg ccc atc tgg aag ttt cca gat gag gag ggc   5403
355 Pro Asp Leu Ser Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly
356                   605                 610                 615
```

VERIFICATION SUMMARY   DATE: 12/30/2005
PATENT APPLICATION: **US/10/560,961**   TIME: 09:43:14

Input Set : **A:\15450PCT00.ST25.txt**
Output Set: **N:\CRF4\12302005\J560961.raw**

L:9 M:270 C: Current Application Number differs, Replaced Current Application No
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date